AppScale ATS on premise

Reference Architecture

This document defines the relevant terms and describes deployment configurations and requirements for AppScale ATS private deployments in on-premise environments.

About

AppScale ATS implements AWS-compatible cloud services over dedicated infrastructure, providing a dedicated private AWS region. ATS enables creation of cost-effective and flexible AWS hybrid cloud environments with a seamless experience for developers and workloads across public and private resources. No special-purpose hardware or unorthodox operating system configurations are required and the entire software stack utilizes open-sourced components.

Subsystems

AppScale ATS deployments consist of a variety of components potentially deployed across many hardware nodes, with each component serving a specific role. It is useful, though, to group components into just four subsystems:

- **Cloud Control (CLC)** - Stateful and stateless components with a cloud-wide mandate.
- **Availability Zone Control (AZC)** - Stateless components responsible for scheduling instances and block storage management in a specific Availability Zone.
- **Storage** - Stateful components implementing backing for object and block storage.
- **Compute** - Stateless components enabling execution of virtual machines.

Logical Configurations

AppScale ATS can be deployed automatically in three standard configurations.

1. **Small (S)** - *One-node deployment.* This is a single-node deployment that cannot be expanded, intended for proof-of-concept setups to demonstrate the features of ATS. Control plane, Storage, and Compute subsystems are all co-located. All features of ATS are available, but performance and capacity of the system are limited. Failure of the node may result in total loss of data. (Use of external persistent volumes would prevent this.)

2. **Medium (M)** - *Hyper-converged deployment (tens of nodes).* Deployment with one or more control-plane nodes (at least one CLC node and one AZC node) and three or more Compute nodes that double as Storage nodes. In this configuration compute and storage capabilities scale in tandem, allowing cost-effective deployments of a range of sizes. There is redundancy for data (block and object) and the possibility of configuring redundancy for outside connectivity. These deployments are limited to one AZ.

3. **Large (L)** - *Dedicated-storage deployment (over forty nodes).* Deployment with two or more control-plane nodes (one CLC node and one or more AZC nodes), three or more dedicated Storage nodes, and one or more dedicated Compute nodes per AZ. In this configuration compute and storage capabilities can be scaled separately, at the additional cost of dedicated Storage nodes. There is redundancy for outside connectivity and for data (block and object).

The three configurations have the following characteristics:

Type	Nodes	Storage	Networking	Console	Use cases
S	1	Linux file system	VPC	yes	POC
M	4 - 40	Ceph on compute	dual-gateway VPC	yes	one-AZ production workloads
L	6 - 100's	Ceph dedicated	dual-gateway VPC	yes	multi-AZ production workloads

Hardware Requirements

Small deployment has all ATS components co-located on a single node. Minimal requirements listed below are adequate for a one-node deployment to function. For production-ready **Medium** and **Large** deployments, nodes should have considerably more capacity, depending on the role (e.g., storage nodes should have enough disk to satisfy EBS and S3 storage requirements multiplied by three for redundancy). Since ATS runs virtual machines using the KVM hypervisor, running ATS wholly in a virtual machine is not recommended.

CPU architecture	x86, with hardware virtualization support
CPU cores	4 minimum, 16+ recommended
RAM	16GB minimum, 32GB+ recommended
Disk	ATS and operating system take up ~5GB, including logs and DB contents, the rest can be allocated for instance-local, EBS, and S3 storage, as needed by the configuration
Network	100Mbps+ Ethernet, with the possibility of configuring multiple NICs to isolate inter-instance network traffic and storage traffic or to create bonded interfaces

The above specifications represent the requirements of ATS components assuming a minimal cloud workload. Optimal hardware configuration depends heavily on workload requirements. Storage-heavy workloads will need significantly more disk storage on the Storage nodes, with SSDs, and possibly with large magnetic spindles storage configured for desired EBS and S3 performance; compute-bound workloads will benefit from more cores and more powerful cores in Compute nodes; and memory-bound workloads will benefit from more RAM with higher data rates in Compute nodes.

Minimum node count requirements for **Medium** and **Large** deployments are as follows:

	Medium	**Large**
CLC and AZC nodes	1+ Controller nodes	2+ Controller nodes
Compute	3+ Compute nodes	1+ Compute nodes
Storage	(co-located with Compute)	3+ Storage nodes

Additional requirements for the environment are listed in the next section.

Deployment Considerations

Beyond hardware specifications, the following needs should be considered:

- Base OS:
  - latest CentOS or RHEL 7.* version
  - SELinux in permissive mode
- Network requirements
  - SSH access to a head-node, password-less SSH between nodes
  - reserved IP address pool, routable from the clients to the head-node, to assign to instances as elastic IP addresses
  - Network switch for communication between nodes (no firewall or limitation of any protocol like IGMP, etc, or any other traffic shaping or manipulation)
  - Optional but highly desirable: a DNS subdomain allocated to each deployment and delegated to the CLC node